



# 70-475<sup>Q&As</sup>

Designing and Implementing Big Data Analytics Solutions

## Pass Microsoft 70-475 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.passapply.com/70-475.html>

100% Passing Guarantee  
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft  
Official Exam Center

-  **Instant Download** After Purchase
-  **100% Money Back** Guarantee
-  **365 Days** Free Update
-  **800,000+** Satisfied Customers





## QUESTION 1

Overview:

Relecloud is a social media company that processes hundreds of millions of social media posts per day and sells advertisements to several hundred companies. Relecloud has a Microsoft SQL Server database named DB1 that stores

information about the advertisers. DB1 is hosted on a Microsoft Azure virtual machine.

Relecloud has two main offices. The offices are located in San Francisco and New York City.

The offices connect to each other by using a site-to-site VPN. Each office connects directly to the Internet.

Relecloud modifies the pricing of its advertisements based on trending topics. Topics are considered to be trending if they generate many mentions in a specific country during a 15-minute time frame. The highest trending topics generate the

highest advertising revenue.

Relecloud wants to deliver reports to the advertisers by using Microsoft Power BI. The reports will provide real-time data on trending topics, current advertising rates, and advertising costs for a given month. Relecloud will analyze the trending

topics data, and then store the data in a new data warehouse for ad-hoc analysis. The data warehouse is expected to grow at a rate of 1 GB per hour or 8.7 terabytes (TB) per year. The data will be retained for five years for the purpose of long-term trending.

Requirements:

Management at Relecloud must be able to view which topics are trending to adjust advertising rates in near real-time.

Relecloud plans to implement a new streaming analytics platform that will report on trending topics.

Relecloud plans to implement a data warehouse named DB2.

Relecloud identifies the following technical requirements:

Social media data must be analyzed to identify trending topics in real-time.

The use of Infrastructure as a Service (IaaS) platforms must be minimized, whenever possible.

The real-time solution used to analyze the social media data must support scaling up and down without service interruption.

Relecloud identifies the following technical requirements for the advertisers:

The advertisers must be able to see only their own data in the Power BI reports.

The advertisers must authenticate to Power BI by using Azure Active Directory (Azure AD) credentials.

The advertisers must be able to leverage existing Transact-SQL language knowledge when developing the real-time streaming solution.



Members of the internal advertising sales team at Relecloud must be able to see only the sales date of the advertisers to which they are assigned.

The internal Relecloud advertising sales team must be prevented from inserting, updating, and deleting rows for the advertisers to which they are not assigned.

The internal Relecloud advertising sales team must be able to use a text file to update the list of advertisers, and then to upload the file to Azure Blob storage.

Relecloud identifies the following requirements for DB1:

Data generated by the streaming analytics platform must be stored in DB1.

The user names of the advertisers must be mapped to CustomerID in a table named Table2.

The advertisers in DB1 must be stored in a table named Table1 and must be refreshed nightly.

The user names of the employees at Relecloud must be mapped to EmployeeID in a table named Table3.

Relecloud identifies the following requirements for DB2:

DB2 must have minimal storage costs.

DB2 must run load processes in parallel.

DB2 must support massive parallel processing.

DB2 must be able to store more than 40 TB of data.

DB2 must support scaling up and down, as required.

Data from DB1 must be archived in DB2 for long-term storage.

All of the reports that are executed from DB2 must use aggregation.

Users must be able to pause DB2 when the data warehouse is not in use.

Users must be able to view previous versions of the data in DB2 by using aggregates.

Relecloud identifies the following requirements for extract, transformation, and load (ETL):

Data movement between DB1 and DB2 must occur each hour.

An email alert must be generated when a failure of any type occurs during ETL processing.

Sample code and data:

You execute the following code for a table named rls\_table1.



```
create function rls_table1 (@CustomerId int, @SalesPersonId int)
    returns table
    with schemabinding
as
return
select 1 as result
from dbo.table1
join dbo.table2 on table1.customerid = Table2.CustomerId
where table2.UserName = suser_sname()
    and table1.customerid = @CustomerId
union all
select 1 as result
from dbo.table1
join dbo.table3 on table1.salespersonid = table3.EmployeeId
where table3.UserName = suser_sname()
    and table1.salespersonid = @SalesPersonId
go
```

You use the following code to create Table1.

```
create table table1 (customerid int, salespersonid int ... ) Go
```

The following is a sample of the streaming data.

User	Country	Topic	Time
user1	USA	Topic1	2017-01-01T00:00:01.0000000Z
user1	USA	Topic3	2017-01-01T00:02:01.0000000Z
user2	Canada	Topic2	2017-01-01T00:01:11.0000000Z
user3	India	Topic1	2017-01-01T00:03:14.0000000Z

You need to implement a solution that meets the data refresh requirement for DB1.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Select and Place:



**Actions**

**Answer Area**

- In DB1, create external objects.
- From the Azure portal, export the storage account key.
- In DB1, create a stored procedure that imports data from an external table to Table1.
- From the Azure portal, create and schedule an Azure Automation job that executes the stored procedure.
- In DB1, create a staging table.



Correct Answer:

**Actions**

**Answer Area**

- In DB1, create external objects.
- 
- 
- 
- In DB1, create a staging table.



- In DB1, create a stored procedure that imports data from an external table to Table1.
- From the Azure portal, export the storage account key.
- From the Azure portal, create and schedule an Azure Automation job that executes the stored procedure.



Azure Data Factory can be used to orchestrate the execution of stored procedures. This allows more complex pipelines to be created and extends Azure Data Factory's ability to leverage the computational power of SQL Data Warehouse.

From scenario:

Relecloud has a Microsoft SQL Server database named DB1 that stores information about the advertisers. DB1 is hosted on a Microsoft Azure virtual machine.

Relecloud identifies the following requirements for DB1:

Data generated by the streaming analytics platform must be stored in DB1.



The advertisers in DB1 must be stored in a table named Table1 and must be refreshed nightly.

Reference: <https://docs.microsoft.com/en-us/azure/machine-learning/machine-learning-data-science-move-sql-server-virtual-machine>

---

## QUESTION 2

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while

the others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You plan to implement a new data warehouse.

You have the following information regarding the data warehouse:

The first data files for the data warehouse will be available in a few days.

Most queries that will be executed against the data warehouse are ad-hoc.

The schemas of data files that will be loaded to the data warehouse change often.

One month after the planned implementation, the data warehouse will contain 15 TB of data.

You need to recommend a database solution to support the planned implementation. Solution: You recommend a Microsoft SQL server on a Microsoft Azure virtual machine. Does this meet the goal?

A. Yes

B. No

Correct Answer: B

---

## QUESTION 3

You need to recommend a platform architecture for a big data solution that meets the following requirements: Supports batch processing Provides a holding area for a 3-petabyte (PB) dataset Minimizes the development effort to implement the solution Provides near real-time relational querying across a multi-terabyte (TB) dataset

Which two platform architectures should you include in the recommendation? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

A. a Microsoft Azure SQL data warehouse

B. a Microsoft Azure HDInsight Hadoop cluster

C. a Microsoft SQL Server database

---



- D. a Microsoft Azure HDInsight Storm cluster
- E. Microsoft Azure Table Storage

Correct Answer: AE

A: Azure SQL Data Warehouse is a SQL-based, fully-managed, petabyte-scale cloud data warehouse. It's highly elastic, and it enables you to set up in minutes and scale capacity in seconds. Scale compute and storage independently, which allows you to burst compute for complex analytical workloads, or scale down your warehouse for archival scenarios, and pay based on what you're using instead of being locked into predefined cluster configurations--and get more cost efficiency versus traditional data warehouse solutions.

E: Use Azure Table storage to store petabytes of semi-structured data and keep costs down. Unlike many data stores--on-premises or cloud-based--Table storage lets you scale up without having to manually shard your dataset. Perform OData-based queries.

Reference: <https://azure.microsoft.com/en-us/services/sql-data-warehouse/> <https://azure.microsoft.com/en-us/services/storage/tables/>

#### QUESTION 4

You have the following Hive query.

```
CREATE TABLE UserVisits (username string, urlvisited string, time date);  
LOAD DATA INPATH `wasb:///Logs\` OVERWRITE INTO TABLE UserVisits;
```

Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the script.

NOTE: Each correct selection is worth one point.

Hot Area:

#### Answer Area

The LOAD DATA statement [answer choice].

inserts data into the UserVisits table record by record
copies the contents of Logs to the UserVisits table directory
moves the contents of Logs to the UserVisits table directory

The UserVisits table type is [answer choice].

dataset
external
manazed

Correct Answer:



### Answer Area

The LOAD DATA statement [answer choice].

inserts data into the UserVisits table record by record
copies the contents of Logs to the UserVisits table directory
moves the contents of Logs to the UserVisits table directory

The UserVisits table type is [answer choice].

dataset
external
manazed

Box 1: moves the contents of Logs to the UserVisits table directory

The LOAD DATA statement streamlines the ETL process for an internal Impala table by moving a data file or all the data files in a directory from an HDFS location into the Impala data directory for that table.

Box 2: Managed

The default kind of table produced by the CREATE TABLE statement is known as an internal table. (Its counterpart is the external table, produced by the CREATE EXTERNAL TABLE syntax.). Internal tables are managed by Impala, and use

directories inside the designated Impala work area.

References:

[https://www.cloudera.com/documentation/enterprise/5-2-x/topics/impala\\_load\\_data.html](https://www.cloudera.com/documentation/enterprise/5-2-x/topics/impala_load_data.html)

[https://www.cloudera.com/documentation/enterprise/5-8-x/topics/impala\\_tables.html#tables](https://www.cloudera.com/documentation/enterprise/5-8-x/topics/impala_tables.html#tables)

### QUESTION 5

You have structured data that resides in Microsoft Azure Blob storage.

You need to perform a rapid interactive analysis of the data and to generate visualizations of the data.

What is the best type of Azure HDInsight cluster to use to achieve the goal? More than one answer choice may achieve the goal. Choose the BEST answer.

- A. Apache Storm
- B. Apache HBase
- C. Apache Hadoop
- D. Apache Spark

Correct Answer: D

A Spark cluster provides in-memory processing, interactive queries, micro-batch stream processing-

Reference: <https://docs.microsoft.com/en-us/azure/hdinsight/hdinsight-hadoop-provision-linux-clusters>





VCE & PDF

PassApply.com

<https://www.passapply.com/70-475.html>

2021 Latest passapply 70-475 PDF and VCE dumps Download

---

[Latest 70-475 Dumps](#)

[70-475 VCE Dumps](#)

[70-475 Practice Test](#)



To Read the [Whole Q&As](#), please purchase the [Complete Version](#) from [Our website](#).

## Try our product !

100% Guaranteed Success  
100% Money Back Guarantee  
365 Days Free Update  
Instant Download After Purchase  
24x7 Customer Support  
Average 99.9% Success Rate  
More than 800,000 Satisfied Customers Worldwide  
Multi-Platform capabilities - [Windows](#), [Mac](#), [Android](#), [iPhone](#), [iPod](#), [iPad](#), [Kindle](#)

We provide exam PDF and VCE of Cisco, Microsoft, IBM, CompTIA, Oracle and other IT Certifications. You can view Vendor list of All Certification Exams offered:

<https://www.passapply.com/allproducts>

## Need Help

Please provide as much detail as possible so we can best assist you.  
To update a previously submitted ticket:



 <p><b>One Year Free Update</b> Free update is available within One Year after your purchase. After One Year, you will get 50% discounts for updating. And we are proud to boast a 24/7 efficient Customer Support system via Email.</p>	 <p><b>Money Back Guarantee</b> To ensure that you are spending on quality products, we provide 100% money back guarantee for 30 days from the date of purchase.</p>	 <p><b>Security &amp; Privacy</b> We respect customer privacy. We use McAfee's security service to provide you with utmost security for your personal information &amp; peace of mind.</p>
---	---	--

Any charges made through this site will appear as Global Simulators Limited.  
All trademarks are the property of their respective owners.  
Copyright © passapply, All Rights Reserved.